

STIC-Biotech/ChemLib

57157

From: STIC-ILL  
Sent: Wednesday, December 26, 2001 6:14 AM  
To: STIC-Biotech/ChemLib  
Subject: RE: 09/497967

-----Original Message-----

From: Fields, Iesha  
Sent: Monday, December 24, 2001 11:02 AM  
T : STIC-ILL  
Subject: 09/497967

Please do a sequence search and interference search on SEQ ID NO:7 for Application 09/497967.

Thank You

Iesha Fields  
Art Unit 1645  
Mailbox 8E-12  
Room 8A-13  
605-1208

Point of Contact:  
Toby Port  
Technical Info. Specialist  
CM1-TEL: 308-3534  
12/14

RECEIVED  
DEC 26 2001  
(STIC)

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 12/26  
Date Completed: 12/26  
Searcher Prep/Review: 10  
Clerical: \_\_\_\_\_  
Online time: 10

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: 1  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST(where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: 8  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

10-20-2011 10:10:10

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 26, 2001, 10:34:04 ; Search time 19.45 Seconds  
(without alignments)  
1832.889 Million cell updates/sec

Title: US-09-497-967-7  
Perfect score: 2540  
Sequence: 1 MKNILVILISLFINQIKS.....QCDNFANFLSLLLSIYLL 468

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	775.5	30.5	395	2 A46031	immobilization sur
2	256.5	10.1	677	2 A42125	trophozoite cystei
3	250	9.8	1766	2 A42125	trophozoite cystei
4	236.5	9.3	667	2 A48579	major surface-labe
5	228.5	9.0	713	2 A35502	cysteine rich prot
6	228	9.0	1274	2 T42017	hypothetical prote
7	218.5	8.6	1372	2 T25933	variant-specific s
8	216.5	8.5	596	2 A45684	variant-specific s
9	197	7.8	557	2 A48434	proteolisin - se
10	191.5	7.5	1297	2 T30274	hypothetical prote
11	190.5	7.5	2823	2 T22759	hypothetical prote
12	188.5	7.4	2823	2 T23064	hypothetical prote
13	188.5	7.4	3102	2 T43291	laminin alpha chai
14	188	7.4	1609	1 MH082	laminin gamma-1 ch
15	188	7.4	5376	2 T42215	laminin gamma-1 ch
16	187.5	7.4	1459	2 T32271	zonadhesin - mouse
17	183.5	7.2	1680	2 A43434	hypothetical prote
18	181.5	7.1	439	2 A36385	furin (EC 3.4.21.7
19	179.5	7.1	1299	2 T43251	surface antigen se
20	179.5	7.1	1548	2 S34583	serine proteinase
21	178.5	7.0	1607	1 MMS52	laminin gamma-1 ch
22	175.5	6.9	3075	2 S14458	laminin gamma-1 ch
23	174.5	6.9	3635	2 T10053	laminin alpha 5 ch
24	173	6.8	3712	2 S18253	laminin alpha-1 ch
25	172	6.8	2718	2 A23475	G surface protein
26	171	6.7	1895	2 T15861	hypothetical prote
27	169	6.7	1827	2 T34288	hypothetical prote
28	167.5	6.6	2704	2 S09118	G surface protein
29	167	6.6	2395	1 S50820	surface protein ty

## ALIGNMENTS

RESULT 1

A46031  
immobilization surface I-antigen precursor - Ichthyophthirius multifiliis (fragment)  
C:Species: Ichthyophthirius multifiliis  
C:Date: 29-Sep-1999 #sequence\_revision 29-Sep-1999 #text\_change 07-Dec-1999  
C:Accession: A46031  
R:Clark, T.G.; McGraw, R.A.; Dickerson, H.W.  
Proc. Natl. Acad. Sci. U.S.A. 89, 6363-6367, 1992  
A:Title: Developmental expression of surface antigen genes in the parasitic ciliate I  
A:Reference number: A46031; MUID:92335298  
A:Accession: A46031  
A:Molecule type: mRNA; protein  
A:Residues: 1-395 <CIA>  
A:Cross-references: GB:M92907; NID:g3628568; PIDN:AAC36158.1; PID:g3628569  
A:Note: the authors translated the codon UUG for residue 330 as Ile  
A:Note: sequence extracted from NCBI backbone (NCBIN:108734, NCBIP:108735); the seque  
C:Genetics:  
A:Genetic code: SGC5  
C:Keywords: glycoprotein; surface antigen  
F:2-395/Product: immobilization surface I-antigen #status experimental <WAT>  
F:156,191,245,281/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 30.5%; Score 775.5; DB 2; Length 395;  
Best Local Similarity 39.8%; Pred. No. 2.8e-43;  
Matches 182; Conservative 36; Mismatches 128; Indels 111; Gaps 18;  
QY 23 CPVGTETNTAGQVD----DLGTPANCVCNCKNFYNNAAAFVPGASTCTPCPOKKDAGAQ 78  
DB 4 CPDGTQIQ-AGLTDVGAADLGT---CVNCRPNFYNGGAA-----OCEAN 44  
QY 79 PNPPATANLVTQCNCVKCPAGTAAGTADYAAITTECVNCRINFYNNENAPNFNAGASTCT 138  
DB 45 GNQDFAAN-----NAARGICV 60  
QY 139 ACPVNRVGGALTAGNATIAVACNVACPTGALDDGVTTDYVRSFTCEVKCRINFYNGN 198  
DB 61 PCQINRVGVTNAGDLATLATQCTQCTGALDDGVTDVDRSAQCVCYKCKPNFYNGG 120  
QY 199 N--GNTP-----FNPG-----KSQCTPCPAIKPANVAQATLGNDATITACQNVACPD 243  
DB 121 SPOGEAPGVGVFAAGAAAGAAVAVTSQCVPCQINK--NDSPATAGAQAANLATQCSNQCT 178  
QY 244 GTTSAAGVNNWVAQTE---CTNCAPNFYNN-----NAPN-----FNPG-----NST 282  
DB 179 GTVLDGVT--LVFNFSATLCVCKRPNFYNGSGSPOGEAPGVGVFAAGAAAGAAVAVTSQ 236  
QY 283 CLCPKANKDYGAETAGGAATLAKQCNIAACPDGTAIASGAT-NVITQTECLNCAANFYF 341  
DB 237 CVPCQLNKN-DSPATAGAQAANLATQCTQCTGALODGVTLVFNFSNSTQCSQCIANYFF 295  
QY 342 DGNFNQAGSSRCACKAPANKVQGVAVATAGGTATLIAQCALECPAGTVLTGTTSTYKQAAS 401



Db 346 -----NAPCTACAGTADKCTKCDANGAAPYLKKTNPSPDPTGTCVSAVDCQG 391  
QY 389 -----TDGTTSTYKQ-----AASECVKCAAN----- 409  
Db 392 SAGYTTSDSDAKCKKCNAPCTACAGTADKCTKCDANGAAPYLKKTNPSPDPTGTCVSA 451  
QY 410 -----FYTQTOTDWMAGIDTCTSCNKLTSGBANLPESAKKNIQCD 451  
Db 452 VDCQGSAGYYT-----DSDVSDAKCKKCNAPCTACA-----GTADKCTKCD 493

RESULT 4  
A48579  
trophozoite surface protein TSP11 - Giardia lamblia  
C:Species: Giardia lamblia  
C>Date: 01-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 29-Jan-1999  
C:Accession: A48579  
R:Ey, P.L.; Khanna, K.K.; Manning, P.A.; Mayrhofer, G.  
Mol. Biochem. Parasitol. 58, 247-257, 1993  
A:Title: A gene encoding a 69-kilodalton major surface protein of Giardia intestinalis  
A:Reference number: A48579; MUID:93241215  
A:Contents: Ad-1  
A:Accession: A48579  
A>Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-667 <EYI>  
A:Cross-references: GB:M95814; NID:gl159106; PID:gl159107  
A:Note: sequence extracted from NCBI backbone (NCBIN:130056, NCBI:P130058)

Query Match 9.3%; Score 236.5; DB 2; Length 667;  
Best Local Similarity 22.7%; Pred. No. 3.4e-08;  
Matches 128; Conservative 49; Mismatches 193; Indels 195; Gaps 30;

QY 7 VILITSLFNQIKSNCPVGTETNTAGQVDDLTGPANVCNQ-----KNFY 52  
Db 1 MLATFYFVISTLTAKCTQTQTQTEAEKCEMVGETEICTRCQTRKVPIDGKCVDATANAN 60  
QY 53 YNNA-----AFVPGASTCTP-----CP-QKKDAGAAPNPANLVTCNV 93  
Db 61 KKNASGDGANQVCKMSVPGTITCTTVSPDGVCSVAANEYFVPPNADATHDSVSCSE 120  
QY 94 KCPAGTA-----TAGGT-----DYAAITEVCNCRINFINENAPNFAGASTC 137  
Db 121 ETPHILANNKQYIGVAGCATCSAPKAPGEDNTPKAACTKCAAGFL--HTP--SEGLSSC 176  
QY 138 -TACPVNRVGGALTAGNAATVAOCNVACPTGTALDDGVTTDYVRSFTECVKRLNFYIN 196  
Db 177 EETCEGYFGHTATRESKKT-----CK-SCTGGSS-----EAPNVKGIQDCLKC-----MYN 222  
QY 197 GNGNTPFPNPKSOCTPCPAIKPANVAQATLGNDA--ITAQCINVA-----CPDG--- 244  
Db 223 EASGNT-----LTCEKCSAQKKPSL-DKTSNCDCTGONCAFCSSSGDCEGDSGFIID 275  
QY 245 -----TISAAGVNNVAQNTCTNCAPNFYNNAPNFNPNSTCLPCPANKDYGA 294  
Db 276 GONCVKSDCKTENCNCACTNPKAAANEVCTECT-----STHHLTPTSCQVQYCOALGNVYA 329  
QY 295 EATAGGAATLAKOCNIA-----CPDGTAIASGATNVILQTECLNCAANFYFDGNNFQAGS 350  
Db 330 GTNADNKA-CKECTVANCKTCD-----QGQCTCNDGDFYKNGDACSPCH 374  
QY 351 SRCKACPANKVOGAVATAGGTATLIAQCALECPAGTVL---TDGTTSTYKQ----- 398  
Db 375 ESKTCSA-----GTA-----SDCT-ECPTGKALKYGNDDTKTCGEGCTTGOGSG 419  
QY 399 -----AASECVK-----AANFYTKQTDWVAGIDTCTSC----- 428  
Db 420 AKTCTGLIDGASTCSCEDTQNEYPQNGICTSTTARTVATCKNSNVANGI--GSSCTNGF 477  
QY 429 -----NKKLTSGAAN 439  
Db 478 LRMMGGCVETTKFPKGSVCBEGANAD 502

RESULT 5  
A35502  
major surface-labeled trophozoite antigen precursor - Giardia lamblia  
C:Species: Giardia lamblia  
C>Date: 09-Nov-1990 #sequence\_revision 09-Nov-1990 #text\_change 09-Sep-1997  
C:Accession: A35502  
R:Gilllin, F.D.; Hagblom, P.; Harwood, J.; Alek, S.B.; Relner, D.S.; McCaffery, M.; So  
Proc. Natl. Acad. Sci. U.S.A. 87, 4463-4467, 1990  
A:Title: Isolation and expression of the gene for a major surface protein of Giardia  
A:Reference number: A35502; MUID:90280395  
A:Accession: A35502  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-713 <GIL>  
A:Cross-references: GB:M33641; NID:gl159131; PID:gl159132  
C:Keywords: surface antigen; transmembrane protein

Query Match 9.0%; Score 228.5; DB 2; Length 713;  
Best Local Similarity 24.4%; Pred. No. 1.2e-07;  
Matches 122; Conservative 34; Mismatches 170; Indels 175; Gaps 33;

QY 23 CPVGTETNTAGQVDDLTGPANVCNQKN-FYIN-----NAAAFVPGASTCTP-----CPQ 71  
Db 69 CP-----QHSAGK-----CTQCGNSPFMYKDGYSSEGLPGHSLCLSSDDGDGVCTE 115  
QY 72 KKDAGAPN-----PPATANLVTCNVKC--PAGTAIAGGATDYAAIITEVCNCRINFINEN 126  
Db 116 -----AAPGFAPVGAAN-TEQSVIACGDTTGVTTIAGGNTYKG-IADCAEC-----S 161  
QY 127 APNFNAGA-----STCTACPVNVRVGGALTAGNAATIVAOCN----- 162  
Db 162 APDATAGAEAGKVATCTKCGVSK-----YLDNVCDVCKAOCNSGNTNKFVAVDDSENGNK 217  
QY 163 VACPT-----GTALDDGVTTDYVRSFTECVKRLNFYNGNNGNTP-----FNP--- 206  
Db 218 VSCSDNLGGVANGDTCYDEQSKIKCTCTDNVYLKTTSEGTSCVQKDOCKDGFPPKD 277  
QY 207 ---GKSQCTPCPAIKP--ANVAQATL--GNDATITACQNVACPDG----- 244  
Db 278 DSSAGNCKLCPNDSTDIANCATCALVSGRSGAALVTCs-ACDGYKPSADKTTCEAVSN 336  
QY 245 --TISAAGVNNVAQNTCTNCAPNFYNNAPNFNPNSTCLPCPA--NKDYGAETAGG 300  
Db 337 CKTPGCKACSNKENEVCVTCDDGSTV-----LTP-TSQCIDSCAKIGNYGA---TEG 386  
QY 301 AATLAKOCNIA---CPDGTAIASGATNVILQTECLNCAANFYFDGNNFQAGSSRCKAC 356  
Db 387 AKLCKECTAANCKTCD-----QGQCOACNDGFKNGDACSPCHESCKTC 432  
QY 357 PANKVOGAVATAGGTATLIAQCALECPAGTVL---TDGTTSTYKQ-----AASECVKCA 407  
Db 433 SA-----GTA-----SDCT-ECPTGKALKYGDGTGKTGCGEGCTTGTGAGACKTCG 477  
QY 408 ANFYTKQTDWVAGIDTCTSC 428  
Db 478 LT-----IDGASYSEC 489

RESULT 6  
T42017  
cysteine rich protein - Giardia intestinalis  
C:Species: Giardia intestinalis  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
C:Accession: T42017  
R:Chen, N.; Upcroft, P.; Upcroft, J.  
Parasitology 111, 423-431, 1995  
A:Title: A Giardia duodenalis gene encoding a protein with multiple repeats of a toxin  
A:Reference number: 22027  
A:Accession: T42017  
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-1274 <CHE>  
A:Cross-references: EMBL:L29079; NID:g951190; PID:g951191; PIDN:AAA74587.1  
A:Experimental source: specific host: Homo sapiens  
C:Superfamily: unassigned ankryrin repeat proteins; ankryrin repeat homology; EGF homology

Query Match 9.08; Score 228; DB 2; Length 1274;  
Best Local Similarity 23.18; Pred. No. 2.2e-07;  
Matches 119; Conservative 35; Mismatches 183; Indels 178; Gaps 30;

QY 21 ANCPV-GTETNAGQVDDLGPANVCNCKNFYNN-----NAAAFPGASTCTPC 69  
DB 801 APCNVEGETCEVGNQA-----QCKTRPGYTINTDKQCKDPEAPCNVEGCETCV-- 852  
QY 70 POKKDAGAP-----NPPATANLVTO-----CNVK-CPAGTAIAGGATDYAAIITEC 115  
DB 853 ----EGNAQOCKTCRPGYTINTDKQCKDPEAPCNVEGCE--TCVEGNA-----QQC 899  
QY 116 VNCNRYNENAPNENAGASTCTACPVNRVGGALTAGNAATIVACNVACPTGTALDDGV 175  
DB 900 KTCRPGY-----TINTDKQCKD-----EAPCNV-----EGC 928  
QY 176 TTDYVRSFTECVKRLNFYNNNG-----NTPFNPKSQCTPCPAIKPANVAQA 225  
DB 929 ETCVEGNAOOCKTCRPGYTINTDKQCKDPEAPCNTP-----NCKTCDNPKTIDNEI-C 981  
QY 226 TLGNDAITTAQCNVACPDGTISAAGV-----NNWVAQNTTECTNCAPNFYNNAPNFPGN 280  
DB 982 TKNDGDYLTPTNQCVPDCT-AISGYDGTDKKCKACNPECAECV-----GFAN 1029  
QY 281 STCLPCPANK--DYGAETAGGAATLAKOCNIA-----CPD-----GTATASGATN--- 324  
DB 1030 NOCTACPVGRKMLQYTNTNTPVNGTCMDQCSYSTNDGCAEGAGIIGTAYCKSKCKNTQQ 1089  
QY 325 -----YVILQTECLNCAANFYF-DG-----NNFQAGSSRC 353  
DB 1090 APLNGCAASSRVAFCATITSGACTKCEGYFLKDGCGYQTDROPFGKQVCSNAQGGNGKC 1149  
QY 354 KACPANKVQAVATAGTATLIAQCALECPAGTVLTDGTTSTYKQAASCVKCAANFYTT 413  
DB 1150 QTC-----ANGLAASDGNCA-ECHS-----TCATCST-ADRAADCKTCATGYIKE 1192  
QY 414 KQTDWVAGIDTCTSCNKKLTSGAEANLPESAKNI 448  
DB 1193 NGDDTTAGL--CKKCEKI-SGCKQCVSSGSSVI 1224

RESULT 7  
T25933  
hypothetical protein W02C12.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T25933  
R:Murray, J.; Wohldmann, P.  
submitted to the EMBL Data Library, December 1996  
A:Description: The sequence of C. elegans cosmid W02C12.  
A:Reference number: Z20112  
A:Accession: T25933  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-1372 <MUR>  
A:Cross-references: EMBL:U08015; PIDN:AA37995.1; GSPDB:GN00022; CESP:W02C12.1  
A:Experimental source: strain Bristol N2; clone W02C12  
C:Genetics:  
A:Gene: CESP:W02C12.1  
A:Map position: 4  
A:Introns: 29/1; 66/1; 774/2; 823/2; 1046/1; 1108/2; 1298/1

Query Match 8.68; Score 218.5; DB 2; Length 1372;  
Best Local Similarity 22.28; Pred. No. 9.7e-07;  
Matches 90; Conservative 35; Mismatches 155; Indels 125; Gaps 20;

QY 79 PN-PPATANLVTOCNVKNCPAGTAIAGGATDYAAIITECVNCRINFYNNENAPNFNAGASTC 137  
DB 905 PNGRPDVGSLKIKEEYLCQAGV-----VRDLCPVCPAGTYH-----SAATGEC 949  
QY 138 TACPYNR-----VSGALTAGNAATIVACNVACPTGTALDDGVTTDVRSFT 184  
DB 950 ELCPIGEYQPLTARTCEFKCAFQGITASGAISEGECKDNCPPGHQY-DSLTS- 1002  
QY 185 ECVKRLNFYNNNGNTNFPNPKSQCTPCPAIKPANVAQATILGNDAITTAQCNVACPDG 244  
DB 1003 -CVTCGYGYQP-----SAGAFECIPC-----GIGKTTLSFATSEDECRCEPDG 1047  
QY 245 -TISAAGV-----NNWVAQNTTECTNCAPNFYNN-----NAPNENPG----- 279  
DB 1048 EQLSASGVCPCQIGTYRSRGENKNCVACPPGTTTATMSTRREQCNTPKCPGQPLVKE 1107  
QY 280 NSTCLPCPANKDYAGATAGGAATLAKQCNIACPDGTATASGATNYVILQTECLN---CA 336  
DB 1108 TKNQCFPRGTFONEQES-----TKLCPADHTTAAPGAT-----AESQCFSTNOCA 1155  
QY 337 ANFYF-----DGNFQAGSSRCACAPANKVOGAVATAGGTATLIAQCALECPAG 385  
DB 1156 TGEYNCNWHANCIDLDPENDVPSEYECRK--PGYRGNTHCTDA-----CNDFC--- 1202  
QY 386 TVLTDGTTSTYKQAASECV-----KCAANFYTTKQTDWVAGI 422  
DB 1203 --LNDGICKNNIGNVEICICKDHFSGRCELRFQASNNKLWIATV 1245

RESULT 8  
A45664  
variant-specific surface protein VSP1267 - Giardia lamblia  
C:Species: Giardia lamblia  
C:Date: 22-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 29-Jan-1999  
C:Accession: A45664  
R:Mowatt, M.R.; Aggarwal, A.; Nash, T.E.  
Mol. Biochem. Parasitol. 49, 215-227, 1991  
A:Title: Carboxy-terminal sequence conservation among variant-specific surface protei  
A:Reference number: A45664; MUID:92131058  
A:Accession: A45664  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-596 <MOW>  
A:Cross-references: GB:M63966; NID:g159140; PID:g159141  
A:Note: sequence extracted from NCBI backbone (NCBIN:77609, NCBIP:77610)

Query Match 8.58; Score 216.5; DB 2; Length 596;  
Best Local Similarity 21.78; Pred. No. 6.1e-07;  
Matches 128; Conservative 45; Mismatches 200; Indels 217; Gaps 33;

QY 5 ILVILIISLFINQIKSANCVPVGTETNTAGQVDDLGTTPANCVCOK----- 49  
DB 4 IAFYLILSTF-----AVDCKNSGNSCEAGQCDTIGDTEICMQCNOGKVPINGICTAHSEE 58  
QY 50 -----NEFYNNAA-----AFVPGASTC 66  
DB 59 AVTNAGCKNKGNTNIEESDKVCGCGNGFYFLHKGCGYKIGEAPGNLICADEASNPART 118  
QY 67 TPCPOKKDAGAPNPPATANLVTOCNVKNCPAGTATAGATDYAAIITECVNCRINFY--- 123  
DB 119 GVCGACKD-GYYKNSDAVAT-ADSCITACEDANCAATCGGAGE-----NCKTCKIDGYFVGA 171  
QY 124 --NE-----NAPFNAGASTCTAC--PVNRVGGALTAGNAATIVAAQC----- 161  
DB 172 TGNEGGCIRKCDATTGNSYKGVAGCAKCEKPRN-----AGPAKCI--ECAAADYLKTEA 222  
QY 162 --NVACPTGTALDDG---VTTDYVRSFTE-CVKCKRLNFYNNNGNTFNPFGKSOCTPC 214  
DB 223 DEQTSVCSEAVCREGKTHEPTTDSAGNKKVCVSC-----GTTNNGGIENGCEC 271  
QY 215 PAIKPANVAQATILGNDAITATA-----QCNVACPDGTISRAGVNNWVA----- 256

```
Db 272 TSRESA-----ARAGTEITCTKCSNNLSPLGDACTCPCAGTAVAGSDSGVCKPCCHNTC 327
QY 257 -----QNTETCNAPNF---YNNAPNPNPNSGNTCL-----PCPANKDY 292
Db 328 AGCQDDRETSCACTACPGYSLLESNG-----ATGRCKVECTGAFITNCADGQCTAN--- 379
QY 293 GAEATAGGAATLAKOCNIACPDG-----TAIASGATNYVILQT---ECLNCAANF-Y 340
Db 380 -----VGG-----AKYCT-CKDGYAPIDGICTAFAAGRDVSVCTATGCKTACTGNYAL 429
QY 341 FDGNFQA-----GSSRCKACAPANKVOGAVAT-AGGTA-----TLIAQCALECPAGVILTD 390
Db 430 LSGGCTNTQTLPGKSVCKAV-ANSNDGCKTCTCANGAOPDPATNFCPLDOSTCAECSTKND 488
QY 391 GTTST-----YKQAASECVCANFYTITKOTDMVAGIDTCTSCNKLTS 435
Db 489 ADACTRCFPGYKTKGNKICIKTESNNGKKID---GIPDCLSCAEPING 535

RESULT 9
A48434
variant-specific surface protein - Giardia lamblia (strain GS/M)
C:Species: Giardia lamblia
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 07-May-1999
C:Accession: A48434
R:Nash, T.E.; Mowatt, M.R.
Mol. Biochem. Parasitol. 51, 219-228, 1992
A:Title: Characterization of a Giardia lamblia variant-specific surface protein (VSP) gene
A:Reference number: A48434; MUID:92244292
A:Accession: A48434
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-557 <NAS>
A:Cross-references: GB:M80480; NID:g159142; PID:g159143

Query Match 7.8%; Score 197; DB 2; Length 557;
Best Local Similarity 20.7%; Pred. No. 1,1e-05;
Matches 110; Conservative 58; Mismatches 21; Indels 152; Gaps 25;

QY 7 VILIISLFTNQIKSANCPCVGTETNTAGQVDDLTGPANCVCOKNFYNN----- 55
Db 1 MFLINCLIASTLAGACSTTOANCAERCEMVGETEICTQCKVVPINGVCEAAASNT.60
QY 56 ----AAAFVPGASTCTPC-----PQKDKAGA----- 77
Db 61 KCKASADEASDQTCGLSTTFMYKGCYDKTGNLRIICKTEGSDAGKCGACNDEKGF 120
QY 78 QPNPPTANLVTQCNVKC--PAGTAIAGGATDYAAIITECVNCRINFYNENAPNPNAGAS 135
Db 121 FDNPDA-ANNVDS-ISCGDATGVIPGSSTKYKGVAGCAKC--TKPSQISENTGTKEA 176
QY 136 TCTACPNRVGALLAGNATIAOCNVA--CPTGTALDDGVTTDYVRSFTECVKRLNF 193
Db 177 TTECNANLYLKAIVSSPTSAT--SCVSAEDCKTYG----PTTDTDSKKKLTCTC---- 225
QY 194 YVNGNGNTFNPGRSQCTPCPAIKPANVAQATL-----GNDATITACQNVACPDGT 245
Db 226 -----STADKGGIDGCSACELLSTTRASTVLISCSACSTNNLSPLKNECMQDCPAGT 278
QY 246 ISAAGV-----NNWVAQNT-CTNCAP-----NFYNNNAPNF-----NPGN 280
Db 279 YADSNVCKPCHTSCASCCKGDNTESSCTACYPGSVLSYGTDTNKTGCIACETGKYLENCA 338
QY 281 SPLCLPCPANKDVGAEATAG-----GAATLAKOCNIACPDGTAIASGATNYVILQTECLNC 335
Db 339 GQCTATIAAGSKYCKSKSGFVNVGCLVSAETARAAPPGST---PDKTNGV-----CTAC 390
QY 336 AANFYFD-GNNFQA-----GSSRCKACAPANK-----VQGAATAGGTATTLIAQCALECPAGT 386
Db 391 TBKYLESSGCGYQAKKFFGNTLCTADAGKCTTCANGQDKDSNGS-----CPA--- 438
```

```
QY 387 VLTDGTTTYYKQAASECVCANF-----TTKQTDWVAGIDTCTSC 428
Db 439 -CPTNCASCADNTKCNKCFSGYVLTADKACKKCGSETSGNIQGVENCISC 488

RESULT 10
T30274
Proteolialisin - sea urchin (Strongylocentrotus purpuratus) (fragment)
C:Species: Strongylocentrotus purpuratus (purple urchin)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C:Accession: T30274
R:Laidlaw, M.; Wessel, G.M.
Development 120, 1325-1333, 1994
A:Title: Cortical granule biogenesis is active throughout oogenesis in sea urchins.
A:Reference number: 220803; MUID:94298531
A:Accession: T30274
A:Status: preliminary; translated from GB/EMBL/DBDB
A:Molecule type: mRNA
A:Residues: 1-1297 <LAI>
A:Cross-references: EMBL:057753; NID:g1373379; PID:g1373380; PIDN:AAB02256.1

Query Match 7.5%; Score 191.5; DB 2; Length 1297;
Best Local Similarity 21.2%; Pred. No. 5.1e-05;
Matches 125; Conservative 63; Mismatches 192; Indels 211; Gaps 37;

QY 21 ANCPVGTETNTAGQVDDLT-----PANCVCOK 49
Db 499 SGCPTTIVTCPAGRI-DCGTNYCVVGARCDGVSDCSNGQDESGCPTTIVTCPAGRIDCGT 557
QY 50 NFYNNAAAFVPGASTCTPCPKKDKAGAPNPATANLVTQCNVKCPAGTAIAGGATDYA 109
Db 558 N--YCVVGARCDGVSDCS--NGQDESGCP-----TIVT-----CPAGRIDCG--TNYC 600
QY 110 AIIITEC---VNCRINFYNENAPNENAGASTCTACPNRV---GGALTAGAATIAOCN- 162
Db 601 VVGARCDGVSDC-----SNGQDEIGCPTTIVTCPAGRVDCGNNYCVVGKCDGVSDCSN 654
QY 163 -----VACPTGTALDDGVTTDYVRSFTECV---KRLNFYNGNNGTTPNPGK 208
Db 655 QODESGCPTTIVTCPPG-RIDCG--TDY-----CVVGARC-----DGVSDCSNGQ 696
QY 209 SQ--CTP---CPA-----IKPANYAQAATLGNDAITTAOCNVACPDGTISAA 249
Db 697 DEICGPTTIVTCPAGRVDCGNNYCVVGKCDGVSDCSNGQDESGCPTTISACPEGRVDCG 756
QY 250 GYNNWVAQNT-CTNCAPNFYNNAPNPNPNSGNTCLPCPANK-DYGAETAGGAATLA 305
Db 757 --NNYCVVGKCDGVSDCSNGQDESGCPTTST---CPEGRVDCGTDYCVGA----- 804
QY 306 KQCNIACPDGTAIASGATNYVILQTECLNCAANFYDGNNFQAGSSRC----- 353
Db 805 -RC-----DGVSDCSNGQDEIGCPTTIVTCPAGRVDCGNNYCVVGKCDGVSDCSNGQDE 858
QY 354 -----KACPANKVQ--GAVATAGGTATLIAQCA-----LECPACTV----- 387
Db 859 SECPPTISACPEGRVDCGNNYCVVGKCDGVSDCSNGQDESGCPTTIVTCPAGRIDCGTN 918
QY 388 -----LTDGTTT-TYKQAASEC-----VKCAANFYTT-KQTDWVAGIDTC--- 425
Db 919 YCVVGARCDGVSDCSNGQDESGCPTTIVTCPAGRVDCGNNYCVVGKCD--GVSDCSNG 975
QY 426 -----TSCNKKLTS---GAEA--NLPESAKNKTQCFANFLSILLI 463
Db 976 QDEBGSFSSCRSGRDC-CEFFTGAEVLESQRVTKLEQLDLLENILVLENTLV 1026

RESULT 11
T22759
hypothetical protein F55H12.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000
C:Accession: T22759
```

A:Experimental source: clone H10E24

R:McMurray, A.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19980

A:Accession: T25096

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-2823 <WIL>

A:Cross-references: EMBL:Z81125; PIDN:CAB03385.1; GSPDB:GN00019; CESP:T22A3.8

A:Experimental source: clone T22A3

C:Genetics:

A:Gene: CESP:T22A3.8

A:Map position: 1

A:Introns: 87/1; 98/1; 126/2; 201/3; 343/3; 406/1; 576/3; 656/1; 825/3; 869/1; 909/1; 96

7/1; 1755/2; 1800/1; 1850/3; 1896/1; 2003/3; 2035/3; 2082/3; 2119/1; 2144/1; 2200/2; 227

C:Superfamily: LDL receptor ligand-binding repeat homology

F:243-279/Domain: LDL receptor ligand-binding repeat homology <LDL>

Query Match 7.5%; Score 190.5; DB 2; Length 2824;  
Best Local Similarity 21.9%; Pred. No. 0.00012;  
Matches 117; Conservative 42; Mismatches 207; Indels 169; Gaps 25;

QY 20 SANCPCVGTET-----NTAGQVDDLTGPANCYCNCKNFYNNAAAFVPGASTCTCPQOKK 74  
DB 2123 SVKCMFGRQAKANNYSTGCDKRLPPTVLRCPKNIRKKTETEDFT----- 2168  
QY 75 AGAQPNNPATANLVTCQNVKCPAGTAAGGATDYAAILTECVNCRINFYNENAPNFENAGA 134  
DB 2169 -----KVWEP-----DEDAFFDNIGVIRIEVNYINGQFGVGI 2202  
QY 135 SICTACPNVRVGGALTAGNAATIAQCNV-----ACPT-----GTALDDGVTTDYVRSFTE 185  
DB 2203 TTV-----RYVGEDAAGNSAECTFDVYIYKSCPSQVYAEGETVLAMQFTTAPTKAE 2256  
QY 186 CVKCRNFY-----YNGNNGNTP-----FNPCKSQCTPCPAIKPANNVAQ 224  
DB 2257 KVKCDNLPTDSRPFMYCDDIMGDYQYGGWSDNTKQIYLPACGQTS--PAVQAINGTV 2314  
QY 225 ATLGNDAIT-----TAQCN--VACP-----DGTISAAG-VNNWVAQNTCC 261  
DB 2315 VSGGQCQIQRURDVIWASADCDRLSCRLMIYPCSDIENGRVSTADELSNIALQYTF 2374  
QY 262 TNCA-----PNFYNNAPNF-----NPGNSTCLPCPANKDYGAETAGGAATLAKQCNIAC 312  
DB 2375 TKNATETIDTTLVNLNQTFTYIRQSDTVDCDP-----SYPIHDTNGNVTICVK-----C 2424  
QY 313 PDGTATASGATNVVILQTECLNCAANFYFDGNNFQAGSRCKACPAKPVQAVATAGTA 372  
DB 2425 PEGT-FANKESN-----KCIDCPINTYRNSTNLD--QLKCTACPGTTVTGDVGTGAVDE- 2474  
QY 373 TLIAQCALECPAGTVLTDG-----TTSTYK-----QAAASECVKCAANFY 411  
DB 2475 ---SQYVNCPIGOFESKGLCNPCDEGTFTGLRKCICCGFDLSTFGGPCIQCPRLGT 2531  
QY 412 TTKQT-----DWVAGIDTCTSCNKKLTSGAEANLPESAKNNICDFANFLSIS 459  
DB 2532 TTSOASTSINSCDTINCIDANTWKNKNTVGPSTPYSEIC---IACEQSTFQNV 2583

RESULT 12

T23064

hypothetical protein T22A3.8 - Caenorhabditis elegans (fragment)

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jul-2000

C:Accession: T23064; T25096

R:Barlow, K.

submitted to the EMBL Data Library, October 1997

A:Reference number: Z19669

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-2823 <WIL>

A:Cross-references: EMBL:AL008585; PIDN:CAA15432.1; GSPDB:GN00019; CESP:T22A3.8

RESULT 13

T43291

laminin alpha chain - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000

C:Accession: T43291

R:Zhu, X.; Rao, G.; Joh, K.; Hall, D.H.; Wadsworth, V.K.; Hutter, H.; Vogel, B.E.; Hu

submitted to the EMBL Data Library, June 1998

A:Description: Expression, function and evolution of laminin alpha chains.

A:Reference number: Z22397

A:Accession: T43291

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-3102 <ZHU>

A:Cross-references: EMBL:AF074902; PIDN:AAC26793.1

C:Genetics:

A:Map position: 1

A>Note: lamal/2

C:Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-lik

A:Experimental source: clone H10E24

R:McMurray, A.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19980

A:Accession: T25096

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-2823 <WIL>

A:Cross-references: EMBL:Z81125; PIDN:CAB03385.1; GSPDB:GN00019; CESP:T22A3.8

A:Experimental source: clone T22A3

C:Genetics:

A:Gene: CESP:T22A3.8

A:Map position: 1

A:Introns: 45/1; 282/2; 312/3; 416/2; 1255/3; 1329/3; 1418/3; 1776/2; 1988/2; 2760/2

C:Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-lik

Query Match 7.4%; Score 188.5; DB 2; Length 2823;  
Best Local Similarity 21.5%; Pred. No. 0.00016;  
Matches 103; Conservative 32; Mismatches 160; Indels 185; Gaps 30;

QY 22 NCPVGTETNTA-----GQVDDL-----TP 41  
DB 860 NCPLYEGNKCEYCSGDFEDPLTGKICETCNGNDIPMGIGNCDSETGKCLKICIGHTTG 919  
QY 42 ANCVNCKNFYNNAAAFVPGASTCTCPQOKKDAGAPNPATANLVTC-----NVKCPAG 98  
DB 920 DSCESC-KEHHWNAQ-----LHTCKPCGCHTOGAVNP-----QCSEENGECECK 963  
QY 99 TATAGGATDYAAILTECVNCRINFYNENAPNFAGASTCTACPNRVGGALTAGNAATIV 158  
DB 964 ENYIG-----AQCDRCK-----ENHGDVENG--CPACDCNDTGS--IGSDCDQVS 1004  
QY 159 AQCNAVCTGTALDDGVTTDYVR-----SFTE--CVKCRNFYNGNNGTFFNPGKSOE- 211  
DB 1005 GQCN--CKQGVF--GKQCDQCRPSYENFTDAGCFQCHNIYGSIEDGKCDQTTGKCECR 1059  
QY 212 -----TPCPAIPKA--NVAQA-----TLGNDAIT-----ITAQCNVACPDDGTISAAGV 251  
DB 1060 ENVEGTMCCKADGYFNITSGDCGDCGDDTGSVDYSCNLTVCQ--VCKPG----- 1110  
QY 252 NNWVAQNTCTNCAPNFYNNAPNFNGNSTCLPCPANKDYGAETAGGAATLAKQCNIA 311  
DB 1111 ---VTGLKCDCLPNFYGLT---SEGCTECPCP-----APGOVCDPIDGSCV 1152  
QY 312 CPDGTATASGATNVVILQTECLNCAANF--YFDGNNFQAGSRCKACPAKPVQAVATAG 369  
DB 1153 CPPNT-----VGEMCENCTTNAWDYHPLNG-----CKLDCSD-----IGSDG 1190  
QY 370 GTA-TLIAQCALECPAGTVLTDGTTSTYKQAAASECVKCAANFYTTKQTDWVAGIDTCTSC 428  
DB 1191 GMCNTFTQCQ--KCKAAV-----GLKCDLCTHGFF-----NFTCEPC 1227







^ ^ ^

Wed Dec 26 12:01:40 2001

us-09-497-967-7.rpr

Page 9

---

